

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2004, 11:28:57 ; Search time 6.03261 Seconds
(without alignments)
319.364 Million cell updates/sec

Title: US-09-657-336A-179
Perfect score: 181
Sequence: 1 NLRLAIEAQCHLLQLTWQIKQLQARILAVERYLKDQ 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 segs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	174	96.1	853	1 ENV_HV1EL	P04581 human immu
2	174	96.1	853	1 ENV_HV1MP	P19551 human immu
3	174	96.1	853	1 ENV_HV1Z2	P12487 human immu
4	174	96.1	855	1 ENV_HV1Z6	P04580 human immu
5	174	96.1	856	1 ENV_HV1B1	P03375 human immu
6	174	96.1	856	1 ENV_HV1H2	P04578 human immu
7	174	96.1	856	1 ENV_HV1H3	P04624 human immu
8	174	96.1	856	1 ENV_HV1H4	P070626 human immu
9	174	96.1	856	1 ENV_HV1PV	P03376 human immu
10	174	96.1	861	1 ENV_HV1BR	P03377 human immu
11	173	95.6	847	1 ENV_HV1S1	P19550 human immu
12	173	95.6	855	1 ENV_HV1OY	P20888 human immu
13	171	94.5	848	1 ENV_HV1JR	P20871 human immu
14	171	94.5	856	1 ENV_HV1MN	P05877 human immu
15	170	93.9	843	1 ENV_HV1Y2	P35961 human immu
16	170	93.9	846	1 ENV_HV1ND	P18799 human immu
17	170	93.9	851	1 ENV_HV1B8	P04582 human immu
18	170	93.9	852	1 ENV_HV1S3	P19549 human immu
19	170	93.9	855	1 ENV_HV1A2	P03378 human immu
20	170	93.9	856	1 ENV_HV1SC	P05878 human immu
21	170	93.9	856	1 ENV_HV1W1	P31872 human immu
22	170	93.9	856	1 ENV_HV1ZH	P05881 human immu
23	170	93.9	861	1 ENV_HV1KB	P31819 human immu
24	170	93.9	865	1 ENV_HV1J3	P04579 human immu
25	170	93.9	867	1 ENV_HV1J3	P12489 human immu
26	170	93.9	868	1 ENV_HV1C4	P05879 human immu
27	169	93.4	859	1 ENV_HV1MA	P04583 human immu
28	167	92.3	847	1 ENV_HV1W2	P05880 human immu
29	165	91.2	847	1 ENV_HV1Z8	P05882 human immu
30	162	89.5	852	1 ENV_HV1BN	P12488 human immu
31	159	87.8	854	1 ENV_SIVCZ	P1281 chimpanzee
32	125	69.1	865	1 ENV_SIVAT	P05886 simian immu
33	125	69.1	877	1 ENV_SIVAG	P27977 simian immu

34	122	67.4	854	1 ENV_SIVAI	Q02837 simian immu
35	116	64.1	768	1 ENV_SIVAI	P27757 simian immu
36	113	62.4	712	1 ENV_HV2S2	P32536 human immu
37	113	62.4	851	1 ENV_HV2D1	P17755 human immu
38	113	62.4	851	1 ENV_HV2G1	P18040 human immu
39	113	62.4	856	1 ENV_HV2NZ	P05883 human immu
40	113	62.4	859	1 ENV_HV2ST	P20872 human immu
41	112	61.9	859	1 ENV_HV2CA	P24105 human immu
42	111	61.3	380	1 ENV_SIVM2	P08810 simian immu
43	111	61.3	858	1 ENV_HV2RO	P04577 human immu
44	111	61.3	885	1 ENV_SIVS4	P12492 simian immu
45	111	61.3	889	1 ENV_SIVSP	P19503 simian immu
46	110	60.8	882	1 ENV_HV2D2	P05885 simian immu
47	109	60.2	859	1 ENV_HV2M1	P15831 human immu
48	108	59.7	880	1 ENV_SIVML	P11267 simian immu
49	107	59.1	821	1 ENV_SIVGB	P22380 simian immu
50	107	59.1	846	1 ENV_SIVMB	P12449 human immu
51	107	59.1	881	1 ENV_SIVMK	P05884 human immu
52	106	58.6	857	1 ENV_HV2KR	P074126 human immu
53	106	58.6	860	1 ENV_HV2BE	P18094 human immu
54	52	28.7	2564	1 SPCQ_HUMAN	Q9h254 homo sapien
55	51.5	28.5	924	1 HXK3_RAT	P27926 ratu
56	50.5	27.9	445	1 EEXL_STAM	Q99TX0 staphylococ
57	50.5	27.9	1411	1 EEAI_HUMAN	Q15075 homo sapien
58	50	27.6	1938	1 MYH4_RABIT	Q28641 ocyctolagus
59	49.5	27.3	583	1 FRIZ_DROVI	P18537 drosophila
60	49.5	27.3	583	1 FRIZ_DROVI	Q24760 drosophila
61	49	27.1	236	1 GT6_SCHMA	P46435 schistosoma
62	49	27.1	1756	1 PEPL_HUMAN	O60437 homo sapien
63	49	27.1	1938	1 MYH4_AEOR	P24733 aequipecten
64	49	27.1	1939	1 MYH4_HUMAN	P74716 fusarium so
65	49	27.1	4349	1 DYHC_FUSSO	Q8bf91 homo sapien
66	49	27.1	8797	1 SNEI_HUMAN	Q8bl66 mus musculu
67	48.5	26.8	790	1 EEAI_MOUSE	Q9x016 thermotoga
68	48	26.5	213	1 OTC_SYNP7	Q935Y4 synechococc
69	48	26.5	305	1 SEHL_HUMAN	Q9h418 homo sapien
70	48	26.5	314	1 YD25_YEAST	Q07657 saccharomyc
71	48	26.5	702	1 AT11_VARAV	P34011 variola vir
72	48	26.5	906	1 CTNI_HUMAN	P35221 homo sapien
73	48	26.5	906	1 CTNI_HUMAN	P26231 mus musculu
74	48	26.5	906	1 CTNI_MOUSE	Q35099 mus musculu
75	48	26.5	1379	1 M3K5_MOUSE	Q9kux2 vibrio chol
76	47.5	26.2	253	1 CYSH_VIBCH	P98203 mus musculu
77	47.5	26.2	962	1 ARVC_HUMAN	Q62245 mus musculu
78	47.5	26.2	969	1 ARVC_MOUSE	Q07889 homo sapien
79	47.5	26.2	1319	1 SOS1_MOUSE	O61308 parascaris
80	47.5	26.2	1333	1 SOS1_HUMAN	P81917 drosophila
81	47.5	26.2	1955	1 PUMA_PARUN	Q50205 mycobacteri
82	47	26.0	376	1 O43A_DROME	Q98f08 rhizobium 1
83	47	26.0	380	1 OXAA_MYCLB	P52630 homo sapien
84	47	26.0	418	1 DADI_RHILLO	Q90339 cyprinus ca
85	47	26.0	851	1 STA2_HUMAN	Q97ej2 clostridium
86	47	26.0	1935	1 MYSS_CYPCA	Q58068 methanococc
87	46.5	25.7	132	1 RS8_GLOAB	P77445 escherichia
88	46.5	25.7	134	1 Y652_METJA	P41793 salmonella
89	46.5	25.7	467	1 EUTE_ECOLI	
90	46.5	25.7	467	1 EUTE_SALTY	

ALIGNMENTS

RESULT 1
ENV_HV1EL
ID ENV_HV1EL STANDARD: PRT: 853 AA.
AC P04581; 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).

OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86245056; PubMed=2424612;
RA Alizon M., Main-Hobson S., Montagnier L., Sonigo P.;
RT "Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
RL Cell 46:63-74(1986).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; K03454; AAA44329.1; -.
DR EMBL; A07108; CA00616.1; -.
DR HIV; K03454; ENV\$SELI.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 31 BY SIMILARITY.
FT CHAIN 32 508 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 509 853 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 206 BY SIMILARITY.
FT DISULFID 125 197 BY SIMILARITY.
FT DISULFID 130 154 BY SIMILARITY.
FT DISULFID 219 248 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 297 330 BY SIMILARITY.
FT DISULFID 376 442 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 853 AA; 96721 MW; F9CD864DAA0D07A5 CRC64;

Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 NLRALIAQOHLQLTWQIKOLARILAVRYLKDQ 37
Db 551 NLRALIAQOHLQLTWQIKOLARILAVRYLKDQ 587

RESULT 2
ENV_HVIMF STANDARD; PRT; 853 AA.
ID ENV_HVIMF
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
GN
OS Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317877; PubMed=1695254;
RA Stevenson M., Haggerty S., Lamontica C., Mann A.M., Meier C.,
RA Wasilak A.;
RT "Cloning and characterization of human immunodeficiency virus type 1
RT variants diminished in the ability to induce syncytium-independent
RT cytolysis.";
RL J. Virol. 64:3792-3803(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; M33943; AAA44850.1; -.
DR PDB; 1A1K; 16-JUN-97.
DR HIV; M33943; ENV\$MFA.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 509
FT CHAIN 510 853 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 203 BY SIMILARITY.
FT DISULFID 126 194 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 216 245 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 294 329 BY SIMILARITY.
FT DISULFID 376 443 BY SIMILARITY.
FT DISULFID 383 416 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).

```
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 3377B993B6F22ABA CRC64;

Query Match 96.1%; Score 174; DB 1; Length 853;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRAIEAQOHLQLTWQIKQARILAVERYLKQ 37
Db 552 NLRAIEAQOHLQLTWQIKQARILAVERYLKQ 588

RESULT 3
ENV_HV122 STANDARD; PRT; 853 AA.
ID ENV_HV122
AC P12487;
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Z2/CDC-Z34 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11683;
RN [1]
RP SEQUENCE FROM N.A.
RA Theodore T., Buckler-White A.;
RL Submitted (NOV-1988) to the HIV data bank.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
FT DISULFID 383 415 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 137 137 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B0B8CBAFF7008 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 853;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRAIEAQOHLQLTWQIKQARILAVERYLKQ 37
Db 551 NLRAIEAQOHLQLTWQIKQARILAVERYLKQ 587

RESULT 4
ENV_HV126 STANDARD; PRT; 855 AA.
ID ENV_HV126
AC P04580;
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., York D., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaraman V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

[illegible][illegible]

```
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 856 AA; 97224 MW; 0BFFB1A18931BB27 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLTRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 37
Db 554 NLTRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 590

RESULT 6
ENV_HV1H2 STANDARD; PRT; 856 AA.
ID ENV_HV1H2 009779;
AC P04578; 13-AUG-1987 (Rel. 05, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11706;
RN [1]
RP MEDLINE=87299196; PubMed=3040055;
RX Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RA "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF033819; AAC82596.1; -
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K33; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENVSHXB2.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
```

```
DR Pfam; PF00516; GP120; 1.
DR AIDS; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT CHAIN 1 30
FT CHAIN 1 30
FT DISULFID 512 511
FT DISULFID 54 856
FT DISULFID 119 74
FT DISULFID 126 205
FT DISULFID 131 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 156 156
FT CARBOHYD 160 160
FT CARBOHYD 186 186
FT CARBOHYD 197 197
FT CARBOHYD 230 230
FT CARBOHYD 234 234
FT CARBOHYD 241 241
FT CARBOHYD 262 262
FT CARBOHYD 276 276
FT CARBOHYD 289 289
FT CARBOHYD 295 295
FT CARBOHYD 301 301
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 356 356
FT CARBOHYD 386 386
FT CARBOHYD 392 392
FT CARBOHYD 397 397
FT CARBOHYD 406 406
FT CARBOHYD 448 448
FT CARBOHYD 463 463
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 624 624
FT CARBOHYD 637 637
FT CARBOHYD 674 674
FT CARBOHYD 750 750
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLTRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 37
Db 554 NLTRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 590

RESULT 7
ENV_HV1H3 STANDARD; PRT; 856 AA.
ID ENV_HV1H3 009779;
AC P04624; 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11707;
RN [1]
RP MEDLINE=87299196; PubMed=3040055;
RX Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RA "Complete nucleotide sequences of functional clones of the AIDS
RT virus.";
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).
RN [2]
RP REVISIONS.
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,
RA Gallo R.C., Wong-Staal F.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; K03455; AAB50262.1; -
DR EMBL; AF038399; AAB99976.1; -
DR EMBL; AF033819; AAC82596.1; -
DR PDB; 1DF4; 26-JAN-00.
DR PDB; 1DF5; 26-JAN-00.
DR PDB; 1DLB; 02-APR-00.
DR PDB; 1G9M; 27-DEC-00.
DR PDB; 1GCI; 19-AUG-98.
DR PDB; 1GZL; 10-OCT-02.
DR PDB; 1K33; 10-OCT-01.
DR PDB; 1K34; 10-OCT-01.
DR HIV; K03455; ENVSHXB2.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
```


RX MEDLINE=85228248; PubMed=2988795;
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
 RA Shaw G.M., Wong-Staal F., Reddy E.P.;
 RT "HIV-III env gene products synthesized in E. coli are recognized by
 RT antibodies present in the sera of AIDS patients.";
 RL Cell 41:979-986(1985).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; M14100; AAA44679.1; -.
 CC PDB; 1JAU; 17-OCT-01.
 CC PDB; 1JAV; 17-OCT-01.
 CC HIV; M14100; ENVSHXB3.
 CC InterPro; IPR000328; Env GP41.
 CC InterPro; IPR000777; GP120.
 CC Pfam; PF00516; GP120; 1.
 CC Pfam; PF00517; GP41; 1.
 CC AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 CC 3D-structure.

FT SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 138 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CARBOHYD 356 356
 FT CARBOHYD 386 386
 FT CARBOHYD 392 392
 FT CARBOHYD 397 397
 FT CARBOHYD 406 406
 FT CARBOHYD 448 448
 FT CARBOHYD 463 463
 FT CARBOHYD 611 611
 FT CARBOHYD 616 616
 FT CARBOHYD 624 624
 FT CARBOHYD 637 637
 FT CARBOHYD 674 674
 FT CARBOHYD 750 750
 FT CARBOHYD 816 816
 SQ SEQUENCE 856 AA; 97188 MW; 3373C68B84C1AFC CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
 Best Local Similarity 97.3%; Pred. No. 2.4e-16;
 Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 1 NLRAIEAQOHLDTWQIKQARILAVERYLKDQ 37
 Db 554 NLRAIEAQOHLDTWQIKQARILAVERYLKDQ 590

RESULT 8

ENV_HVILW STANDARD; PRT; 856 AA.
 AC 070626;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_TaxID=82834;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95127297; PubMed=7826699;
 RA Reitz M.S. Jr., Hall L., Robert-Guy M., Lautenberger J., Hahn B.M.,
 RA Shaw G.M., Kong L.I., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
 RT "Viral variability and serum antibody response in a laboratory worker
 RT infected with HIV type 1 (HIV type IIB).";
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U12055; AAA76690.1; -.
 CC PDB; 1IF3; 02-MAY-01.
 CC GlycoSuiteDB; Q70626; -.
 CC InterPro; IPR000328; Env GP41.
 CC InterPro; IPR000777; GP120.
 CC Pfam; PF00516; GP120; 1.
 CC Pfam; PF00517; GP41; 1.
 CC AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 CC 3D-structure.
 FT SIGNAL 1 30
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 289
 FT CARBOHYD 295 295
 FT CARBOHYD 301 301
 FT CARBOHYD 332 332
 FT CARBOHYD 339 339
 FT CHAIN 31 511
 FT CHAIN 512 856
 FT DISULFID 54 74
 FT DISULFID 119 205
 FT DISULFID 126 196
 FT DISULFID 131 157
 FT DISULFID 218 247
 FT DISULFID 228 239
 FT DISULFID 296 331
 FT DISULFID 378 445
 FT DISULFID 385 418
 FT CARBOHYD 88 88
 FT CARBOHYD 136 136
 FT CARBOHYD 141 141
 FT CARBOHYD 156 156
 FT CARBOHYD 160 160
 FT CARBOHYD 186 186
 FT CARBOHYD 197 197
 FT CARBOHYD 230 230
 FT CARBOHYD 234 234
 FT CARBOHYD 241 241
 FT CARBOHYD 262 262
 FT CARBOHYD 276 276
 FT CARBOHYD 289 28

```
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 96938 MW; 0C241332CF7E6687 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 37
Db 554 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 590

RESULT 9
ENV_HV1PV STANDARD; PRT; 856 AA.
ID ENV_HV1PV
AC P03376;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511157; PubMed=2982104;
RA Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
RT AIDS/lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: K02083; AAB59873.1; -
DR EMBL: X01762; CAA25903.1; ALT_SEQ.
DR PIR: A03974; VCLJVL.
DR HIV: K02083; ENVSPV22.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
```

```
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97339 MW; 5FCD81DC3C1209B3 CRC64;

Query Match 96.1%; Score 174; DB 1; Length 856;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 37
Db 554 NLRAIEAQOHLQLTWQIKQLQARILAVERYLKDQ 590

RESULT 10
ENV_HV1BR STANDARD; PRT; 861 AA.
ID ENV_HV1BR
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099333; PubMed=2981635;
RA Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;
RT "Nucleotide sequence of the AIDS virus, LAV.";
RL Cell 40:9-17(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; K02013; AAB59751.1; -.
DR EMBL; A04321; CA00352.1; -.
DR PIR; A03975; VCLJLV.
DR PDB; 1ERF; 20-FEB-02.
DR HIV; K02013; ENV$BRU.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 30
FT CHAIN 31 516
FT CHAIN 517 861
FT DISULFID 54 74
FT DISULFID 119 210
FT DISULFID 126 201
FT DISULFID 131 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 336
FT DISULFID 383 450
FT DISULFID 390 423
FT CARBOHYD 88 88
FT CARBOHYD 136 136
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 235 235
FT CARBOHYD 239 239
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 337 337
FT CARBOHYD 344 344
FT CARBOHYD 361 361
FT CARBOHYD 391 391
FT CARBOHYD 397 397
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 453 453
FT CARBOHYD 468 468
FT CARBOHYD 616 616
FT CARBOHYD 621 621
FT CARBOHYD 630 630
FT CARBOHYD 642 642
FT CARBOHYD 679 679
FT CARBOHYD 755 755
FT CARBOHYD 821 821
SQ SEQUENCE 861 AA; 97487 MW; 04DE2BAD4E4FD63A CRC64;

Query Match 96.1%; Score 174; DB 1; Length 861;
Best Local Similarity 97.3%; Pred. No. 2.4e-16;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 NLRAIEAOHLQLTWQIKQLQARILAVERYLKQD 37
Db 559 NLRAIEAOHLQLTWQIKQLQARILAVERYLKQD 595
```

```
RESULT 11
ENV_HV1S1
ID ENV_HV1S1 STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
```

```
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11691;
RN [1]
RP MEDLINE=90347835; PubMed=2384920;
RX Cheng-Mayer C., Quiroga M., Tung J.W., Levy J.;
RT "Viral determinants of human immunodeficiency virus type 1 T-cell or
RT macrophage tropism, cytopathogenicity, and CD4 antigen modulation.";
RL J. Virol. 64:4390-4398(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65024; AAA45072.1; -.
DR PDB; 1OBE; 15-MAY-97.
DR HIV; M38428; ENV$SF162.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 502
FT CHAIN 503 847
FT DISULFID 53 73
FT DISULFID 118 203
FT DISULFID 125 194
FT DISULFID 130 155
FT DISULFID 216 245
FT DISULFID 226 237
FT DISULFID 294 328
FT DISULFID 374 435
FT DISULFID 381 408
FT CARBOHYD 87 87
FT CARBOHYD 135 135
FT CARBOHYD 154 154
FT CARBOHYD 186 186
FT CARBOHYD 195 195
FT CARBOHYD 232 232
FT CARBOHYD 239 239
FT CARBOHYD 260 260
FT CARBOHYD 274 274
FT CARBOHYD 293 293
FT CARBOHYD 299 299
FT CARBOHYD 329 329
FT CARBOHYD 336 336
FT CARBOHYD 352 352
FT CARBOHYD 382 382
FT CARBOHYD 388 388
FT CARBOHYD 392 392
FT CARBOHYD 398 398
FT CARBOHYD 401 401
FT CARBOHYD 438 438
FT CARBOHYD 454 454
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7FE2AB CRC64;

Query Match 95.6%; Score 173; DB 1; Length 847;
Best Local Similarity 94.6%; Pred. No. 3.3e-16;
```


[illegible]

FT	CARBOHYD	336	336	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	359	359	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	389	389	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	395	395	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	399	399	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	405	405	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	458	458	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	610	610	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	615	615	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	624	624	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	636	636	N-LINKED (GLCNAC. .)	(POTENTIAL).					
FT	CARBOHYD	815	815	N-LINKED (GLCNAC. .)	(POTENTIAL).					
SQ	SEQUENCE	855 AA;	97476 MW; 9CF82A607ADD62DA CRC64;							
Query Match		95.6%;	Score 173;	DB 1;	Length 855;					
Best Local Similarity		94.6%;	Pred. No. 3.3e-16;							
Matches	35;	Conservative	1;	Mismatches	1;					
				Indels	0;					
				Gaps	0;					
QY	1 NLRAIEAQCHLLQTWGIKQLQARILAVERYLKDD 37 : : : : :									
Dd	553 NLRAIEAQCHLLQTWGIKQLQARVLAVERYLKDQ 589									
RESULT 13										
ID	ENV_HVIJUR	STANDARD;	PRT;	848 AA.						
AC	P20871;									
DT	01-FEB-1991 (Rel. 17, Created)									
DT	01-FEB-1991 (Rel. 17, Last sequence update)									
DT	10-OCT-2003 (Rel. 42, Last annotation update)									
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].									
GN	ENV.									
OS	Human immunodeficiency virus type 1 (JRCSF isolate) (HIV-1).									
OX	Viruses; Retroviral viruses; Retroviridae; Lentivirus.									
NCBI_TaxID=11688;	[1]									
RN	SEQUENCE FROM N.A.									
RA	Koyanagi S., Chen I.S.Y.;									
RL	Submitted (DEC-1988) to the HIV data bank.									

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).										

CC	EMBL, M38429; AAB03749.1; -. PDB; 1CE4; 1B-MAR-99.									
DR	HIV; M38429; ENV\$JRCSF.									
DR	InterPro; IPR000328; Env_GP41.									
DR	InterPro; IPR000777; GP120.									
DR	Pfam; PF00516; GP120; 1.									
DR	Pfam; PF00517; GP41; 1.									
KM	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal; 3D-structure.									
KW	3D-structure.									
FT	SIGNAL	1	32							
FT	CHAIN	33	503		EXTERIOR MEMBRANE GLYCOPROTEIN.					
FT	CHAIN	504	848		TRANSMEMBRANE GLYCOPROTEIN.					
FT	DISULFID	53	73		BY SIMILARITY.					
FT	DISULFID	118	203		BY SIMILARITY.					
FT	DISULFID	125	194		BY SIMILARITY.					
FT	DISULFID	130	154		BY SIMILARITY.					
FT	DISULFID	216	245		BY SIMILARITY.					
FT	DISULFID	226	237		BY SIMILARITY.					
FT	DISULFID	294	328		BY SIMILARITY.					
FT	DISULFID	374	437		BY SIMILARITY.					
FT	DISULFID	381	410		BY SIMILARITY.					
FT	CARBOHYD	87	87		N-LINKED (GLCNAC. .)					
FT	CARBOHYD	134	134		N-LINKED (GLCNAC. .)					
FT	CARBOHYD	137	137		N-LINKED (GLCNAC. .)					

```

FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 617 617 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 629 629 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 808 808 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 848 AA; 96475 MW; 20767F51227EC3F3 CRC64;

```

```

Query Match 94.5%; Score 171; DB 1; Length 848;
Best Local Similarity 91.9%; Pred. No. 6.2e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NLRAIEAQHLLQLTWQIKQLARILAVERYLKDQ 37
Db 546 NLRAIEAQHMLQLTWGIKQLQARVLAVERYLKDQ 582

```

```

RESULT 14
ENV_HV1MN STANDARD; PRT; 856 AA.
AC P05877;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farello K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL virology 164:531-536(1988).
CC -I- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M17449; AAA44857.1; -
CC PDB; 1ACY; 31-JUL-94.
CC PDB; 1F58; 29-DEC-99.
CC PDB; 1NIZ; 25-FEB-03.
CC PDB; 1NJ0; 25-FEB-03.
CC HIV; M17449; ENV$MN.

```

```

DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL 1 29
FT CHAIN 30 513
FT CHAIN 514 856
FT DISULFID 53 73
FT DISULFID 118 210
FT DISULFID 125 201
FT DISULFID 130 162
FT DISULFID 223 252
FT DISULFID 233 244
FT DISULFID 301 335
FT DISULFID 381 445
FT DISULFID 388 418
FT CARBOHYD 87 87
FT CARBOHYD 129 129
FT CARBOHYD 135 135
FT CARBOHYD 140 140
FT CARBOHYD 141 141
FT CARBOHYD 146 146
FT CARBOHYD 161 161
FT CARBOHYD 165 165
FT CARBOHYD 191 191
FT CARBOHYD 202 202
FT CARBOHYD 246 246
FT CARBOHYD 267 267
FT CARBOHYD 281 281
FT CARBOHYD 294 294
FT CARBOHYD 300 300
FT CARBOHYD 336 336
FT CARBOHYD 343 343
FT CARBOHYD 359 359
FT CARBOHYD 365 365
FT CARBOHYD 395 395
FT CARBOHYD 401 401
FT CARBOHYD 405 405
FT CARBOHYD 406 406
FT CARBOHYD 413 413
FT CARBOHYD 448 448
FT CARBOHYD 465 465
FT CARBOHYD 612 612
FT CARBOHYD 617 617
FT CARBOHYD 626 626
FT CARBOHYD 638 638
FT CARBOHYD 816 816
SQ SEQUENCE 856 AA; 97140 MW; D197D809940BE732 CRC64;

```

```

Query Match 94.5%; Score 171; DB 1; Length 856;
Best Local Similarity 91.9%; Pred. No. 6.2e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 NLRAIEAQHLLQLTWQIKQLARILAVERYLKDQ 37
Db 555 NLRAIEAQHMLQLTWGIKQLQARVLAVERYLKDQ 591

```

```

RESULT 15
ENV_HV1Y2 STANDARD; PRT; 843 AA.
AC P35961;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (YU-2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=36377;

```

```
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93021387; PubMed=1404605;
RA Li Y., Hui H., Burgess C.J., Price R.W., Sharp P.M., Hahn B.H.,
RA Shaw G.M.;
RT "Complete nucleotide sequence, genome organization, and biological
RT properties of human immunodeficiency virus type 1 in vivo: evidence
RT for limited defectiveness and complementation.";
RL J. Virol. 66:6587-6600(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M93258; -, NOT_ANNOTATED_CDS.
CC PIR; H44001; H44001.
CC PDB; 1G9N; 27-DEC-00.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane; Signal;
CC 3D-structure.
KW SIGNAL
FT 1 29
FT CHAIN 30 489 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSMEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DFD971C918B71 CRC64;
```

Query Match 93.9%; Score 170; DB 1; Length 843;
Best Local Similarity 91.9%; Pred. No. 8.5e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 NLRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 37
Db 541 NLRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 577
RESULT 16
ENV_HVIND STANDARD; PRT; 846 AA.
AC P18799;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
RA Hampe A., Chermann J.C.;
RT "Nucleotide sequence of HIV-1-NDK: a highly cytopathic strain of the
RT human immunodeficiency virus.";
RL Gene 81:275-284(1989).
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
CC AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M27323; AAA44873.1; -.
CC PIR; JQ0066; VCLJND.
CC HIV; M27323; ENV$NDK.
CC InterPro; IPR000328; Env GP41.
CC InterPro; IPR000777; GP120.
CC Pfam; PF00516; GP120; 1.
CC Pfam; PF00517; GP41; 1.
CC AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;
CC Signal.
KW SIGNAL
FT 1 29
FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 502 846 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 200 BY SIMILARITY.
FT DISULFID 125 191 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 213 242 BY SIMILARITY.
FT DISULFID 223 234 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```

FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 96476 MW; 8A3B9DA527DE2E83 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 846;
Best Local Similarity 91.9%; Pred. No. 8.5e-16;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLLRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 37
Db 544 NLLRAIEAQHLLQLTWGIGIKQLQARVLAVERYLRDQ 580

RESULT 17
ENV_HV1B8 STANDARD; PRT; 851 AA.
AC P04582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11684;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Joseph S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K02011; AAA44661.1; -.
DR PDB; 1DDH; 13-JAN-99.
DR PDB; 1HHG; 31-OCT-93.
DR PDB; 1QO3; 02-JAN-00.
DR PDB; 1S2T; 24-DEC-97.
DR HIV; K02011; ENV5B8.
DR GlycoSuiteDB; P04582; -.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
KW 3D-structure.
FT SIGNAL. 1 30
FT CHAIN 31 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 507 851 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.

```

```

FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 440 BY SIMILARITY.
FT DISULFID 385 413 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 851;
Best Local Similarity 94.6%; Pred. No. 8.6e-16;
Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLLRAIEAQHLLQLTWQIKQLQARILAVERYLKDQ 37
Db 549 NLLRAIEAQHLLQLTWGIGIKQLQARVLAVERYLRDQ 585

RESULT 18
ENV_HV1S3 STANDARD; PRT; 852 AA.
AC P19549;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SF33 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90317906; PubMed=2370688;
RA York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;
RT "Human immunodeficiency virus type 1 cellular host range,
RT replication, and cytopathicity are linked to the envelope region of
RT the viral genome.";
RL J. Virol. 64:4016-4020(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch.

CC -----
 DR EMBL; AY352275; AAQ17031.1; -.
 DR PDB; 1MEQ; 11-DEC-02.
 DR HIV; M38427; ENV\$SF33.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal;
 3D-structure.

FT SIGNAL 1 31 BY SIMILARITY.
 FT CHAIN 32 506 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 507 852 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.
 FT DISULFID 118 206 BY SIMILARITY.
 FT DISULFID 125 197 BY SIMILARITY.
 FT DISULFID 130 156 BY SIMILARITY.
 FT DISULFID 219 248 BY SIMILARITY.
 FT DISULFID 229 240 BY SIMILARITY.
 FT DISULFID 297 331 BY SIMILARITY.
 FT DISULFID 377 439 BY SIMILARITY.
 FT DISULFID 384 412 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 852 AA; 96663 MW; EE7BHF8D23C910D CRC64;

Query Match 93.9%; Score 170; DB 1; Length 852;
 Best Local Similarity 91.9%; Pred. No. 8.6e-16;
 Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NLRRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 37
 Db 550 NLRRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 586

RESULT 19
 ENV_HV1A2 STANDARD; PRT; 855 AA.
 AC P03378;

DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.

OS Human immunodeficiency virus type 1 (ARV2/SF2 isolate) (HIV-1).
 OC Viruses; Retrovirus; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11685;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=8509453; PubMed=2578227;
 RA Sanchez-Pescador R., Power M.D., Barr P.J., Steimer K.S.,
 RA Stempelen M.M., Brown-Shimer S.L., Gee W.W., Renard A., Randolph A.,
 RA Levy J.A., Dina D., Luciw P.A.;
 RT "Nucleotide sequence and expression of an AIDS-associated retrovirus
 (ARV-2).";
 RL Science 227:484-492(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; K02007; AAB59882.1; -.
 DR PIR; A03976; VCLJ42.
 DR HIV; K02007; ENV\$SF2.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
 KW Signal.

FT SIGNAL 1 29
 FT CHAIN 30 509
 FT CHAIN 510 855
 FT DISULFID 53 73
 FT DISULFID 118 208
 FT DISULFID 125 199
 FT DISULFID 130 155
 FT DISULFID 221 250
 FT DISULFID 231 242
 FT DISULFID 299 333
 FT DISULFID 380 442
 FT DISULFID 387 415
 FT CARBOHYD 87 87
 FT CARBOHYD 129 129
 FT CARBOHYD 140 140
 FT CARBOHYD 154 154
 FT CARBOHYD 158 158
 FT CARBOHYD 184 184
 FT CARBOHYD 190 190
 FT CARBOHYD 200 200
 FT CARBOHYD 233 233
 FT CARBOHYD 244 244
 FT CARBOHYD 265 265
 FT CARBOHYD 279 279
 FT CARBOHYD 292 292
 FT CARBOHYD 298 298
 FT CARBOHYD 304 304
 FT CARBOHYD 334 334
 FT CARBOHYD 341 341
 FT CARBOHYD 358 358
 FT CARBOHYD 364 364
 FT CARBOHYD 388 388
 FT CARBOHYD 394 394
 FT CARBOHYD 400 400
 FT CARBOHYD 408 408
 FT CARBOHYD 445 445
 FT CARBOHYD 458 458
 FT CARBOHYD 461 461
 FT CARBOHYD 610 610
 FT CARBOHYD 615 615
 FT CARBOHYD 624 624
 FT CARBOHYD 636 636
 FT CARBOHYD 815 815


```

SQ      SEQUENCE      855 AA;   97438 MW;   A3BC20573AAC41A2 CRC64;

Query Match          93.9%; Score 170; DB 1; Length 855;
Best Local Similarity 91.9%; Pred. No. 8.6e-16;
Matches    34; Conservative    2; Mismatches    1; Indels     0; Gaps     0;

Oy       1 NLRAIEAQOHLLQLTWQIKQLARILAVERYLKDQ 37
Db        553 NLRRAIEAQOHLLQLTWGWIKQLARVLAVERYLKDQ 589

RESULT 20
ENV_HV1SC STANDARD; PRT; 856 AA.
AC P05878;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (SC isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reltz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates.";
RL Virology 164:531-536(1988).
CC -I- MISCELLANEOUS: THE SC ISOLATE WAS TAKEN FROM AN ARC PATIENT IN
CC 1984 IN SOUTHERN CALIFORNIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR DR EMBL; M17450; -, NOT_ANNOTATED_CDS.
DR DR HIV; M17450; ENVSSC.
DR DR InterPro; IPR000328; Env_GP41.
DR DR InterPro; IPR000777; GP120.
DR DR Pfam; PF00516; GP120; 1.
DR DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT FT SIGNAL . 1 29
FT FT CHAIN 30 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT FT CHAIN 511 856 TRANSMEMBRANE GLYCOPROTEIN.
FT FT SITE 760 760 IN-FRAME TERMINATION CODON.
FT FT DISULFID 53 73 BY SIMILARITY.
FT FT DISULFID 118 206 BY SIMILARITY.
FT FT DISULFID 125 197 BY SIMILARITY.
FT FT DISULFID 130 160 BY SIMILARITY.
FT FT DISULFID 219 247 BY SIMILARITY.
FT FT DISULFID 228 239 BY SIMILARITY.
FT FT DISULFID 296 330 BY SIMILARITY.
FT FT DISULFID 376 439 BY SIMILARITY.
FT FT DISULFID 383 412 BY SIMILARITY.
FT FT CARBOHYD 87 87 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 129 129 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 135 135 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 140 140 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 143 143 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 159 159 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 163 163 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 189 189 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 198 198 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT FT CARBOHYD 234 234 N-LINKED (GLCNAC. . . ) (POTENTIAL).

```

Query Match	Best Local Similarity	Score	DB	Length
Matches 34; Conservative	91.9%;	170;	Pred. No. 8.6e-16;	856;
		2; Mismatches	1; Indels	0; Gaps
Qy	1	NLRAIEAQOHLLQTLTWQIKOLQARILAVERYLKDQ	37	
Db	554	NLRAIEAQOHLLQTLTWGIKOLQARILAVERYLRDQ	590	
ENV_HV1W1	STDARD;	PRT;	856 AA.	
AC	P31872;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
OS	ENV.			
OS	Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=31678;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RK	MEDLINE=86218077; PubMed=2423250;			
RA	Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F., "Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS."			
RT	AIDS.			
RL	Cell 45:637-648(1986).			
CC	-1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.			
CC	PIR; A24774; VCLJ3W.			
DR	PDB; 1LB0; 04-DEC-02.			
DR	PDB; 1LCX; 04-DEC-02.			
DR	InterPro; IPR000328; Env GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal; 3D-structure.			
FT	SIGNAL	1	29	
FT	CHAIN	30	510	
FT	CHAIN	511	856	
FT	DISULFID	53	73	
FT	DISULFID	118	205	
FT	DISULFID	125	196	
FT	DISULFID	125	196	

```

FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 330 BY SIMILARITY.
FT DISULFID 376 444 BY SIMILARITY.
FT DISULFID 383 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97526 MW; DB68D1E49C404DE9 CRC64;

```

Query Match 93.9%; Score 170; DB 1; Length 856;
 Best Local Similarity 91.9%; Pred. No. 8.6e-16;
 Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLIRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 37
 Db 554 NLIRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 590

```

RESULT 22
ENV_HV1ZH STANDARD; PRT; 856 AA.
AC P05881;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire H2321 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OK NCBI_TaxID=11692;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89228766; Pubmed=2713163;
RA Strinivasan A., York D., Butler D. Jr., Jannoun-Nasr R., Getchell J.,
  McCormick J., Ou C.Y., Myers G., Smith T., Chen E.;
RT "Molecular characterization of HIV-1 isolated from a serum collected
  in 1976: nucleotide sequence comparison to recent isolates and
  generation of hybrid HIV.";
RL AIDS Res. Hum. Retroviruses 5:121-129(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL; M15896; AAB53948.1; -.
DR PIR; A44963; A44963.
DR HIV; M15896; ENV$Z321.
DR InterPro; IPR000328; Env GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
DR AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
  signal.
KW Signal.
FT CHAIN 1 29
FT CHAIN 30 511
FT CHAIN 512 856
FT DISULFID 53 73
FT DISULFID 118 206
FT DISULFID 125 197
FT DISULFID 130 153
FT DISULFID 219 248
FT DISULFID 229 240
FT DISULFID 297 331
FT DISULFID 379 445
FT DISULFID 386 418
FT CARBOHYD 87 87
FT CARBOHYD 132 132
FT CARBOHYD 138 138
FT CARBOHYD 152 152
FT CARBOHYD 156 156
FT CARBOHYD 183 183
FT CARBOHYD 198 198
FT CARBOHYD 242 242
FT CARBOHYD 263 263
FT CARBOHYD 277 277
FT CARBOHYD 294 294
FT CARBOHYD 302 302
FT CARBOHYD 332 332
FT CARBOHYD 339 339
FT CARBOHYD 355 355
FT CARBOHYD 364 364
FT CARBOHYD 387 387
FT CARBOHYD 393 393
FT CARBOHYD 398 398
FT CARBOHYD 402 402
FT CARBOHYD 411 411
FT CARBOHYD 448 448
FT CARBOHYD 461 461
FT CARBOHYD 462 462
FT CARBOHYD 465 465
FT CARBOHYD 611 611
FT CARBOHYD 616 616
FT CARBOHYD 637 637
SQ SEQUENCE 856 AA; 96909 MW; 8396B3F8BBD174E CRC64;

```

Query Match 93.9%; Score 170; DB 1; Length 856;
 Best Local Similarity 94.6%; Pred. No. 8.6e-16;
 Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLIRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 37
 Db 554 NLIRAIEAQOHLQLTWQIKQLARILAVERYLKDQ 590

```

RESULT 23
ENV_HV1KB STANDARD; PRT; 861 AA.
ID ENV_HV1KB
AC P31819;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
  glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```


FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 302 302 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 455 455 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 825 825 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 865 AA; 97809 MW; 28828BC2314ADCAC CRC64;

Query Match 93.9%; Score 170; DB 1; Length 865;
 Best Local Similarity 91.9%; Pred. No. 8.7e-16;
 Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NLRRAIEAQOHLTLFTWQIKQLQARILAVERYLKDQ 37
 |||||
 Db 563 NLRRAIEAQOHLTLFTWQIKQLQARILAVERYLKDQ 599

RESULT 25
 ID ENV_HV1J3 STANDARD; PRT; 867 AA.
 AC P12489;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (JH3 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11694;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89352108; PubMed=2669897;
 RA Komiyama N., Hattori N., Inoue J., Sakuma S., Kurimura T., Yoshida M.;
 RT "Nucleotide sequences of gag and env genes of a Japanese isolate of HIV-1 and their expression in bacteria."
 RL AIDS Res. Hum. Retroviruses 5:411-419(1989).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M21138; AAB03526.1; -.
 DR HIV; M21138; ENV5JH3.
 DR InterPro; IPR000328; Env_GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120; 1.
 DR Pfam; PF00517; GP41; 1.
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.
 KM SIGNAL.
 FT SIGNAL 1 30
 FT CHAIN 31 516 EXTERIOR MEMBRANE GLYCOPROTEIN.
 FT CHAIN 517 867 TRANSMEMBRANE GLYCOPROTEIN.
 FT DISULFID 53 73 BY SIMILARITY.

FT DISULFID 118 217 BY SIMILARITY.
 FT DISULFID 125 208 BY SIMILARITY.
 FT DISULFID 130 160 BY SIMILARITY.
 FT DISULFID 230 259 BY SIMILARITY.
 FT DISULFID 240 251 BY SIMILARITY.
 FT DISULFID 308 342 BY SIMILARITY.
 FT DISULFID 388 457 BY SIMILARITY.
 FT DISULFID 395 430 BY SIMILARITY.
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 188 188 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 253 253 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 408 408 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 867 AA; 98399 MW; 5F2310146B8E8680 CRC64;

Query Match 93.9%; Score 170; DB 1; Length 867;
 Best Local Similarity 94.6%; Pred. No. 8.7e-16;
 Matches 35; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NLRRAIEAQOHLTLFTWQIKQLQARILAVERYLKDQ 37
 |||||
 Db 565 NLRRAIEAQOHLTLFTWQIKQLQARILAVERYLKDQ 601

RESULT 26
 ID ENV_HV1C4 STANDARD; PRT; 868 AA.
 AC P05879;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
 GN ENV.
 OS Human immunodeficiency virus type 1 (CDC-451 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11687;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87041461; PubMed=3490666;
 RA Desai S.M., Kalyanaraman V.S., Casey J.M., Srinivasan A.,
 RA Andersen P.R., Devare S.G.;
 RT "Molecular cloning and primary nucleotide sequence analysis of a distinct human immunodeficiency virus isolate reveal significant divergence in its genomic sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:8380-8384(1986).
 RN [2]
 RP SEQUENCE OF 34-43.

Query Match	Best Local Similarity	Score 170; DB 1; Length 868;
Matches 35; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	93.9%; Pred. No. 8.7e-16;	
RA MEDLINE=90253924; PubMed=2187500;		
RA Kalyanaraman V.S. Rodriguez V., Veronese F., Rahman R., Lusso P.,		
RA Devico A.L., Copeland T., Oroszian S., Gallo R.C., Sarngadharan M.G.;		
RT "Characterization of the secreted, native gp120 and gp160 of the human		
RT immunodeficiency virus type 1.";		
RL AIDS Res. Hum. Retroviruses 6:371-380(1990).		
CC -----		
CC This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC the European Bioinformatics Institute. There are no restrictions on its		
CC use by non-profit institutions as long as its content is in no way		
CC modified and this statement is not removed. Usage by and for commercial		
CC entities requires a license agreement (See http://www.isb-sib.ch/announce		
CC or send an email to license@isb-sib.ch).		
CC -----		
DR EMBL; M13137; AAA44311.1; -.		
DR PIR; C25523; VCLJH4.		
DR HIV; M13137; ENVSCDC45.		
DR InterPro; IPR000328; Env GP41.		
DR InterPro; IPR000777; GP120.		
DR Pfam; PF00516; GP120, 1.		
DR Pfam; PF00517; GP41, 1.		
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;		
KW Signal.		
FT SIGNAL 1 33		
FT CHAIN 34 522	EXTERIOR MEMBRANE GLYCOPROTEIN (GP120).	
FT CHAIN 523 868	TRANSMEMBRANE GLYCOPROTEIN.	
FT DISULFID 55 75	BY SIMILARITY.	
FT DISULFID 120 216	BY SIMILARITY.	
FT DISULFID 127 207	BY SIMILARITY.	
FT DISULFID 132 163	BY SIMILARITY.	
FT DISULFID 229 258	BY SIMILARITY.	
FT DISULFID 239 250	BY SIMILARITY.	
FT DISULFID 307 341	BY SIMILARITY.	
FT DISULFID 387 456	BY SIMILARITY.	
FT DISULFID 394 429	BY SIMILARITY.	
FT CARBOHYD 89 89	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 131 131	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 138 138	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 139 139	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 142 142	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 162 162	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 166 166	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 195 195	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 198 198	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 208 208	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 245 245	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 252 252	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 273 273	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 287 287	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 300 300	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 306 306	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 312 312	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 342 342	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 349 349	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 365 365	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 371 371	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 395 395	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 405 405	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 409 409	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 459 459	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 473 473	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 623 623	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 628 628	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 637 637	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 649 649	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT CARBOHYD 828 828	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SO SEQUENCE 868 AA; 98698 MW; A11527FC52A6F0C8 CRC64;		

[illegible]


```

FT CARBOHYD 364 364 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD- 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 639 639 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 819 819 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 859 AA; 97109 MW; DBCF9AA52E3ABF29 CRC64;

Query Match 93.4%; Score 169; DB 1; Length 859;
Best Local Similarity 91.9%; Pred. No. 1.2e-15;
Matches 34; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLRRAIEAQGHLLQLTWQIKQLQARILAVERYLKDQ 37
Db 556 NLRRAIEAQGHLLQLTWQIKQLQARVLAVERYLQDQ 592

RESULT 28
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT at risk for AIDS.";
RL Science 232:1548-1553(1986).
CC -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M12507; AAB12990.1; -.
CC DR HIV; M12507; ENV$WMJ2.
CC DR InterPro; IPR000328; Env GP41.
CC DR InterPro; IPR000777; GP120.
CC DR Pfam; PF00516; GP120; 1.
CC DR Pfam; PF00517; GP41; 1.
CC KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
CC Signal.
CC FT SIGNAL. 1 29
CC FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.
CC FT CHAIN 502 847 TRANSMEMBRANE GLYCOPROTEIN.
CC FT DISULFID 53 73 BY SIMILARITY.
CC FT DISULFID 118 202 BY SIMILARITY.
CC FT DISULFID 125 193 BY SIMILARITY.
CC FT DISULFID 130 152 BY SIMILARITY.
CC FT DISULFID 215 244 BY SIMILARITY.
CC FT DISULFID 225 236 BY SIMILARITY.

```

```
CC FT DISULFID      293          326    BY SIMILARITY.  
CC FT DISULFID     372         435    BY SIMILARITY.  
FT DISULFID       379        408    BY SIMILARITY.  
FT CARBOHYD       87         87  
FT CARBOHYD      134        134    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      140        140    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      151        151    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      155        155    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      183        183    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      184        184    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      194        194    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      231        231    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      238        238    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      259        259    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      273        273    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      286        286    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      292        292    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      327        327    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      334        334    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      350        350    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      356        356    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      380        380    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      386        386    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      390        390    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      400        400    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      438        438    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      450        450    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      602        602    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      607        607    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      616        616    N-LINKED (GLCNAC . .) (POTENTIAL).  
FT CARBOHYD      628        628    N-LINKED (GLCNAC . .) (POTENTIAL).  
SQ SEQUENCE       847 AA;   96466 MW; CDIE33D73AASBCAE CRC64;  
  
Query Match           92.3%; Score 167; DB 1; Length 847;  
Best local Similarity 89.2%; Pred. No. 2.2e-15;  
Matches 33; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY              1 NLRLRAIAEQOHLQLTWGRIKOLQRILAVERYLKDQ 37  
Db             545 NLRLRAIDAQQHLLQLTWTWGIRKLQLRVLAVERYLARDQ 581  
               |||||:|||||:|||||:|||||:|||||::||  
  
RESULT 29  
ENV_HVI1Z8  
ID ENV_HVI1Z8 STANDARD; PRT; 863 AA.  
AC P05882;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI_TaxId=11681;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88281278; PubMed=3395517;  
RA Yourho J., Josephs S.F., Reitz M.S. Jr., Zagury D., Wong-Staal F., Gallo R.C.;  
RT "Nucleotide sequence analysis of the env gene of a new Zairian isolate of HIV-1.";  
RL AIDS Res. Hum. Retroviruses 4:165-173(1988).  
CC -I- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD ZAIREAN MALE.
```

CC	EMBL; J03653; AAA44684.1; -.
DR	HIV; J03653; ENV\$JY1.
DR	InterPro; IPR000328; Env GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane; Signal.
FT	SIGNAL. 1 29
FT	CHAIN 30 518
FT	CHAIN 519 863
FT	DISULFID 53 73
FT	DISULFID 118 217
FT	DISULFID 125 208
FT	DISULFID 130 160
FT	DISULFID 230 259
FT	DISULFID 240 251
FT	DISULFID 308 342
FT	DISULFID 388 452
FT	DISULFID 395 425
FT	CARBOHYD 87 87
FT	CARBOHYD 129 129
FT	CARBOHYD 136 136
FT	CARBOHYD 142 142
FT	CARBOHYD 143 143
FT	CARBOHYD 159 159
FT	CARBOHYD 163 163
FT	CARBOHYD 194 194
FT	CARBOHYD 199 199
FT	CARBOHYD 209 209
FT	CARBOHYD 246 246
FT	CARBOHYD 274 274
FT	CARBOHYD 288 288
FT	CARBOHYD 301 301
FT	CARBOHYD 307 307
FT	CARBOHYD 343 343
FT	CARBOHYD 350 350
FT	CARBOHYD 365 365
FT	CARBOHYD 396 396
FT	CARBOHYD 402 402
FT	CARBOHYD 406 406
FT	CARBOHYD 412 412
FT	CARBOHYD 455 455
FT	CARBOHYD 468 468
FT	CARBOHYD 469 469
FT	CARBOHYD 472 472
FT	CARBOHYD 618 618
FT	CARBOHYD 623 623
FT	CARBOHYD 632 632
FT	CARBOHYD 644 644
FT	CARBOHYD 823 823
SEQ	SEQUENCE 863 AA; 97743 MW; B729CBA6FAD1641 CRC64;
Query Match	91.2%; Score 165; DB 1; Length 863;
Best local Similarity	89.2%; Pred. No. 4.3e-15;
Matches 33; Conservative	2; Mismatches 2; Indels 0; Gaps 0;
Oy	1 NLIRAEIAQOHLQLTWQIKOLARILAVERYLKDQ 37 : :
Db	561 NLIIRAIEAQOHMLQLTWGIKOLARVLAVESYLKDQ 597
RESULT 30	
ENV_HV1BN	STANDARD; PRT; 852 AA.
AC	P12488;
DT	01-OCT-1989 (Rel. 12, Created)
DT	01-OCT-1989 (Rel. 12, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Envelope polypeptide GP160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN	ENV.

[illegible]

FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

Query Match 89.5%; Score 162; DB 1; Length 852;
Best Local Similarity 86.5%; Pred. No. 1.1e-14;
Matches 32; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NLRAIEAQOHLQLTWQIKOLQARILAVERYLKDQ 37
Db 550 NLMAIEAQOHLQLTWGIGIKOLQARVLAVERYLKDQ 586

Search completed: June 2, 2004, 11:42:57
Job time : 7.03261 secs

THIS PAGE BLANK (USPTO)